Transcriptome analysis reveals determinant stages controlling human embryonic stem cell commitment to neuronal cells

Received for publication, May 12, 2017, and in revised form, September 17, 2017 Published, Papers in Press, September 26, 2017, DOI 10.1074/jbc.M117.796383

Yuanyuan Li[‡1], Ran Wang[‡1], Nan Qiao[§], Guangdun Peng[‡], Ke Zhang[‡], Ke Tang[¶], Jing-Dong J. Han[§], and Naihe Jing[‡‖2]

From the [‡]State Key Laboratory of Cell Biology, CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology and [§]Chinese Academy of Sciences Key Laboratory of Computational Biology, Chinese Academy of Sciences-Max Planck Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yue Yang Road, Shanghai 200031, the [¶]Institute of Life Science, Nanchang University, Nanchang, Jiangxi 330031, and the [‖]School of Life Science and Technology, ShanghaiTech University, 100 Haike Road, Shanghai 201210, China

Edited by Xiao-Fan Wang

Proper neural commitment is essential for ensuring the appropriate development of the human brain and for preventing neurodevelopmental diseases such as autism spectrum disorders, schizophrenia, and intellectual disorders. However, the molecular mechanisms underlying the neural commitment in humans remain elusive. Here, we report the establishment of a neural differentiation system based on human embryonic stem cells (hESCs) and on comprehensive RNA sequencing analysis of transcriptome dynamics during early hESC differentiation. Using weighted gene co-expression network analysis, we reveal that the hESC neurodevelopmental trajectory has five stages: pluripotency (day 0); differentiation initiation (days 2, 4, and 6); neural commitment (days 8–10); neural progenitor cell proliferation (days 12, 14, and 16); and neuronal differentiation (days 18, 20, and 22). These stages were characterized by unique module genes, which may recapitulate the early human cortical development. Moreover, a comparison of our RNA-sequencing data with several other transcriptome profiling datasets from mice and humans indicated that Module 3 associated with the day 8–10 stage is a critical window of fate switch from the pluripotency to the neural lineage. Interestingly, at this stage, no key extrinsic signals were activated. In contrast, using CRISPR/Cas9–mediated gene knockouts, we also found that intrinsic hub transcription factors, including the schizophrenia-associated *SIX3* gene and septo-optic dysplasia-related *HESX1* gene, are required to program hESC neural determination. Our results improve the understanding of the mechanism of neural commitment in the human brain and may help elucidate the etiology of human mental disorders and advance therapies for managing these conditions.

Brain development is one of the most complicated and hierarchical events in mammals. A series of temporal processes, including neural commitment, patterning and sub-regionalization, neurogenesis, and neuronal network formation, are required to generate a functional brain (1–4). Defects in any of these processes will lead to neurodevelopmental diseases such as autism spectrum disorders, schizophrenia, depression, and epilepsy. Neural commitment occurs after gastrulation and initiates the program to form neural epithelium (5–7). Many studies on amphibians and rodents reveal that the inhibition of BMP[3] signaling is an evolutionally conserved mechanism for neural induction; in addition, intrinsic factors play pivotal roles to direct fate transition from pluripotency to neural lineage (8–11). Nonetheless, how the neural commitment is ensured in human brain has not been clearly elucidated. Systematically understanding the critical role of signaling pathways and transcriptional factors in modulating and shaping human brain development is hindered by limited accessibility to early embryos and inadequate amounts of stage-specific and cell type–specific materials. These problems may now be solved by the use of human embryonic stem cells (hESCs). Establishment of *in vitro* differentiation models that recapitulate normal development will facilitate the study in brain development and neurological disorders.

The establishment of neural differentiation protocols for hESCs makes it possible to investigate early events, including neural commitment in humans (12–15). hESCs exhibit the restricted capacity to generate various subtypes of functional neurons by responding to extrinsic signals (16–19), which reca-

This work was supported in part by the "Strategic Priority Research Program" of the Chinese Academy of Sciences Grant XDA01010201, National Key Basic Research and Development Program of China Grants 2014CB964804 and 2015CB964500, and National Natural Science Foundation of China Grants 31430058, 31571513, 31630043, 91519314, and 31661143042. The authors declare that they have no conflicts of interest with the contents of this article.

This article was selected as one of our Editors' Picks.

RNA-seq data have been deposited at Gene Expression Omnibus under accession number GSE 103715.

This article contains supplemental Figs. S1–S7 and Tables S1–S3.

[1] Both authors contributed equally to this work.

[2] To whom correspondence should be addressed: Tel.: 86-21-54921381; Fax: 86-21-54921011; E-mail: njing@sibcb.ac.cn.

[3] The abbreviations used are: BMP, bone morphogenetic protein; RNA-Seq, RNA sequencing; hESC, human embryonic stem cell; NPC, neural progenitor; qPCR, quantitative PCR; EB, embryonic body; aEB, attached EB; FPKM, fragment per kilobase per million; WGCNA, weighted gene co-expression network analysis; TF, transcription factor; CSI, Connection Specificity Index; FDR, false discovery rate; hg, human genome; PCA, principal component analysis; PCC, Pearson correlation coefficient; PC, principal component.

pitulate brain development *in vivo*. Global gene expression profiling technologies such as microarray and RNA sequencing (RNA-Seq) enable highly sensitive analysis of transcriptome associated with the neurodevelopment of hESCs. By using RNA-Seq data with nine samples from days 0 to 77 of the differentiating hESCs, van de Leemput *et al.* (20) establish a CORTECON system to study human cerebral cortex development *in vitro*. Meanwhile, single-cell RNA-Seq is applied to identify various classes of neural progenitors (NPCs) and neurons, which is helpful for mapping the early human brain cells (21). Microarray analysis shows that the non-canonical WNT signaling pathway is important for partitioning neural *versus* epidermal fate during neural induction (22). It has been shown that the early neurodevelopment of hESCs advances much quicker than that *in vivo* (13, 15, 23). Therefore, the insufficient representation of differentiating time points analyzed by RNA-Seq or the low resolution of the microarray technique limits the outcome of systematic analysis on fast and transient cell fate transition such as neural induction.

In this study, we adapted and developed an *in vitro* hESC neural differentiation system, ending up with a high percentage of dorsal forebrain neurons. By specific co-expression gene assays of transcriptome data with 12 samples prepared every other day between differentiation day 0 and day 22, we show that the following five distinct stages exist during the early neural differentiation of hESCs: pluripotency (day 0); differentiation initiation (day 2/4/6); neural commitment (day 8/10); NPC proliferation (day 12/14/16); and neuronal differentiation stage (day 18/20/22). Expression profiling comparison of gene modules and transcription factor (TF) gene groups among several systems reveals that the Module 3-associated day 8/10 stage is a critical window for the fate transition from the pluripotency to the neural epithelium. Moreover, *PAX6*, *SIX3*, *SIX6*, *HESX1,* and *ID3* are identified as key hub TF genes of this stage. The loss-of-function of either the *SIX3* or *HESX1* gene, mediated by the CRISPR/Cas9 gene editing system, leads to compromised neural commitment of hESCs.

Results

Directed differentiation of hESCs mimics the early cortical development in vivo

To investigate the regulatory mechanisms of human neural commitment, we first adapted the previous protocols (12) and standardized an *in vitro* hESC (H9 line) neural differentiation system, with EB formation for 6 days, attached EB (aEB) for 10 days, sphere in N2 for 6 days, and then single cells replated in N2B27 for 4 weeks (Fig. 1*A*). Twelve RNA samples were collected every other day from day 0 to day 22 and subjected to RNA sequencing. Quantitative real-time PCR (RT-qPCR) assays with 12 RNA samples were performed to briefly characterize the differentiation process of hESCs. Following the differentiation, the expression of pluripotent genes *POU5F1*, *SOX2*, *NANOG* was decreased, and the expression of neuroectoderm genes *POU3F1* and *ZNF521*, as well as neural epithelial marker genes such as *PAX6* and *SOX1,* was increased and reached the peak at day 12. The expression of anterior forebrain progenitor marker genes *FOXG1*, *OTX2*, and *SIX3* was up-regulated at around day 16, followed by the elevation of neuronal marker genes *TUJ1* (*TUBB3*) and *MAP2* around days 16–22 (Fig. 1*B*). The similar expression pattern could be obtained with another hESC line H1 (supplemental Fig. S1*A*), and the correlation coefficient between H9 and H1 cells is high ($R^2 = 0.95$, supplemental Fig. S1*B*). To further confirm the synchronization of the neural differentiation protocol, we randomly collected 36 single cells at days 0, 4, 8, 14, and 20, respectively, and performed single cell qPCR for *OCT6* (*POU3F1*), *SIX3*, and *SOX1* genes. The results show that the majority of single cells show the comparable expression level for each gene, and the expression pattern of these genes is similar to the results from population cell samples (supplemental Fig. S1*D*). Consistent with RNA-expressing patterns, immunostaining assays revealed that many cells in day 10 aEB were double-positive for the NPC markers PAX6 and NESTIN (62.5%, Fig. 1*C* and supplemental Fig. S1*C*); in addition, PAX6 and OTX2 were co-expressed in those NPCs, indicating a cerebral cortical identity (71.3%, Fig. 1*C* supplemental Fig. S1*C*). Numerous NPCs in day 14 aEBs were PAX6 and SOX1 double-positive (85.1%, Fig. 1*C* supplemental Fig. S1*C*). At day 30, the majority of the cells not only displayed the prolonged neurites but also was positive for mature neuron markers MAP2 and NEUN (82.8%, Fig. 1*C* supplemental Fig. S1*C*). At differentiation day 50, the predominant neuronal subtype was TBR1+ (78.8%) and VGLUT1/2+ (88.6%) positive cortical glutamatergic projection neurons (Fig. 1, *D* and *E*). There were few TH+ dopaminergic neurons (5.1%) but no VAChT+ and GAD67+ neurons (Fig. 1, *D* and *E*).

The results above demonstrate that our protocol can successfully program hESCs to differentiate into cortical NPCs and then cortical projection neurons, and this procedure is reliable and could be applied to different hESC lines.

Transcriptome analysis reveals five possible sub-stages of the hESC neural differentiation

Next, RNA-Seq assays with 12 RNA samples were conducted. Approximately 30 million sequencing reads of every sample were mapped to the human genome hg19. The average number of detectable genes was about 15,000 in each sample with fragments per kilobase per million (FPKM) more than 0.1 in at least one sample (supplemental Fig. S2, *A* and *B*). Correlation analysis confirmed that the normalized RNA-Seq tag counts of marker genes were consistent with their expression levels assessed by RT-qPCR (Fig. 2*A*). Unsupervised hierarchical clustering analysis revealed that there were three main stages during the early hESC neural differentiation, designed as EB (days 0, 2, 4, and), aEB (days 8, 10, 12, 14, and 16), and Sphere (days 18, 20, and 22) stage, respectively (Fig. 2*B*). Within the EB stage, days 2, 4, and 6 were closely clustered, compared with day 0. In the aEB stage, there were two sub-clusters, one with days 8 and 10 and the other with days 12, 14, and 16 (supplemental Fig. S2*C*). Principal component analysis (PCA) revealed that the three major stages could be further divided into five sub-groups (Fig. 2*C* and supplemental Fig. S2*D*). PC1 accounted for the temporal trajectory of differentiation, and PC2 separated the differentiation stages in more detail. As distributed on the transition line from the PC1-negative direction to the PC1-positive direction, day 8/10 probably is a critical period

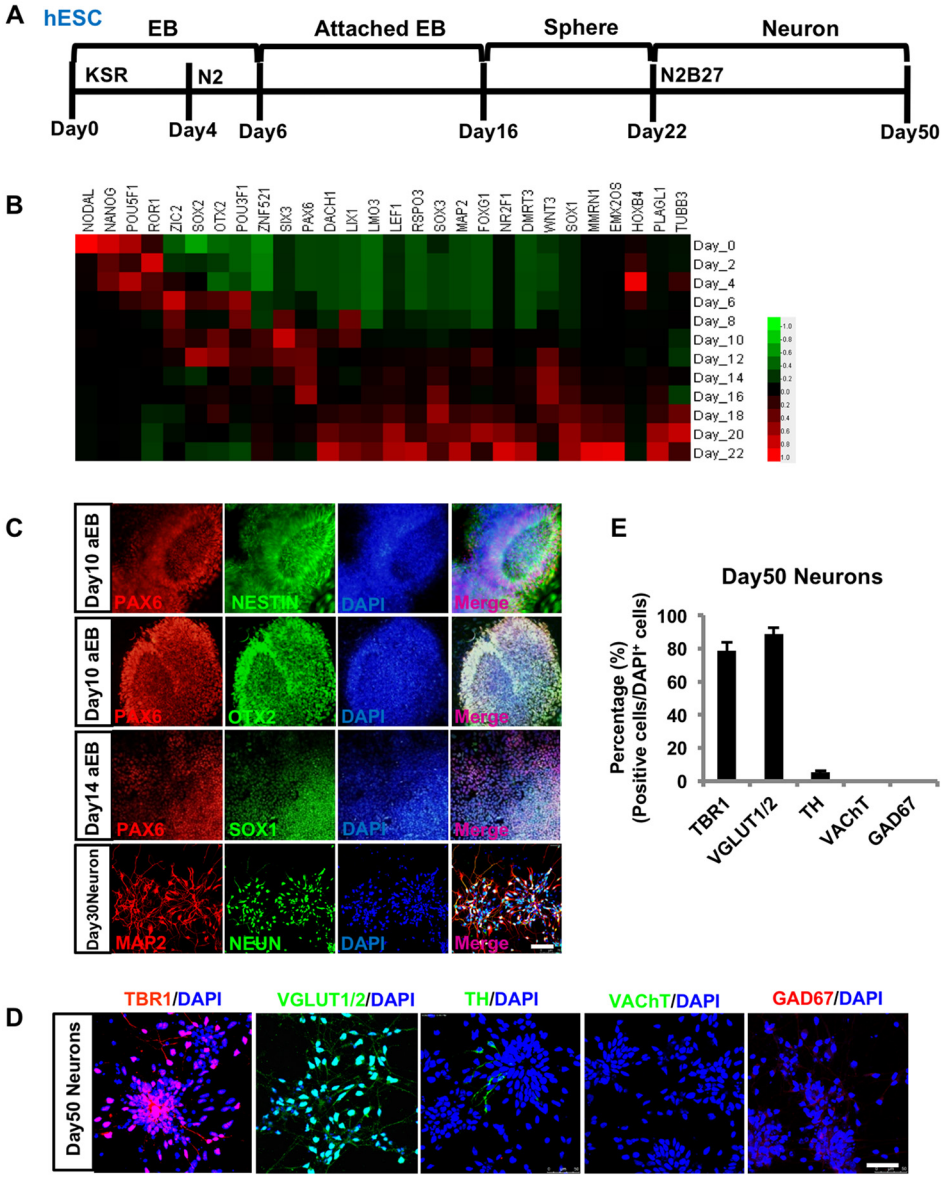


Figure 1. Efficient neural differentiation from human embryonic stem cells. *A,* schematic representation of the hESC neural differentiation method over 50 days. *B,* gene expression heat map of RT-qPCR results for different maker genes at the time indicated. *C,* double immunocytochemistry analysis of PAX6 with NESTIN and OTX2, respectively, in human attached EBs at day 10 (*top*). Double immunocytochemistry analysis of PAX6 with SOX1 in human attached EB at day 14, and MAP2 with NEUN in human neurons at day 30 (*bottom*). *D,* immunofluorescence analysis of TBR1, VGLUT1/2, TH, VAChT, and GAD67 with DAPI in human neurons at day 50. *E,* quantification of data in *D*. *Scale bars,* 50 μm (*C* and *D*). To quantify the differentiation efficiency, three to five fields were randomly selected, and all the experiments were performed in triplicate.

of fate transition. Consistently, clustering analysis using the highest PC-loading genes in the PC1 (24) (top 100 PC1-negative and -positive genes) suggested that day 8 in cluster 3 is a critical fate transition period during hESC neural differentiation (supplemental Fig. S2*E*). Especially, functional enrichment analysis showed that the genes with positive PC loadings in PC1 showed enrichment for forebrain development, neuron development, and axonogenesis. In contrast, the genes with negative PC loadings in PC1 were associated with RNA metabolic process, apoptosis, and cell cycle (Fig. 2*D*). Taken together, the dynamic gene expression profiles revealed a temporal developmental trajectory of hESC neural differentiation, and day 8/10 of cluster 3 is a critical fate transition period.

Gene-network modules define five temporal sub-stages of hESC neural differentiation

To systematically investigate the correlation of gene expression profiles, we performed weighted gene co-expression network analysis (WGCNA) to identify distinct co-expression modules among the clusters of associated transcripts without supervision and bias (25–27). We first focused on modules that can distinguish five sub-stages during the hESC neural differentiation. Among dozens of modules, five modules, marked in *cyan, dark orange, light yellow, brown,* and *steel blue*, were tightly associated with these five sub-stages, respectively (Fig. 3*A*). The expression of genes identified in each module was specifically enriched in the corresponding sub-stage (Fig. 3*B*).

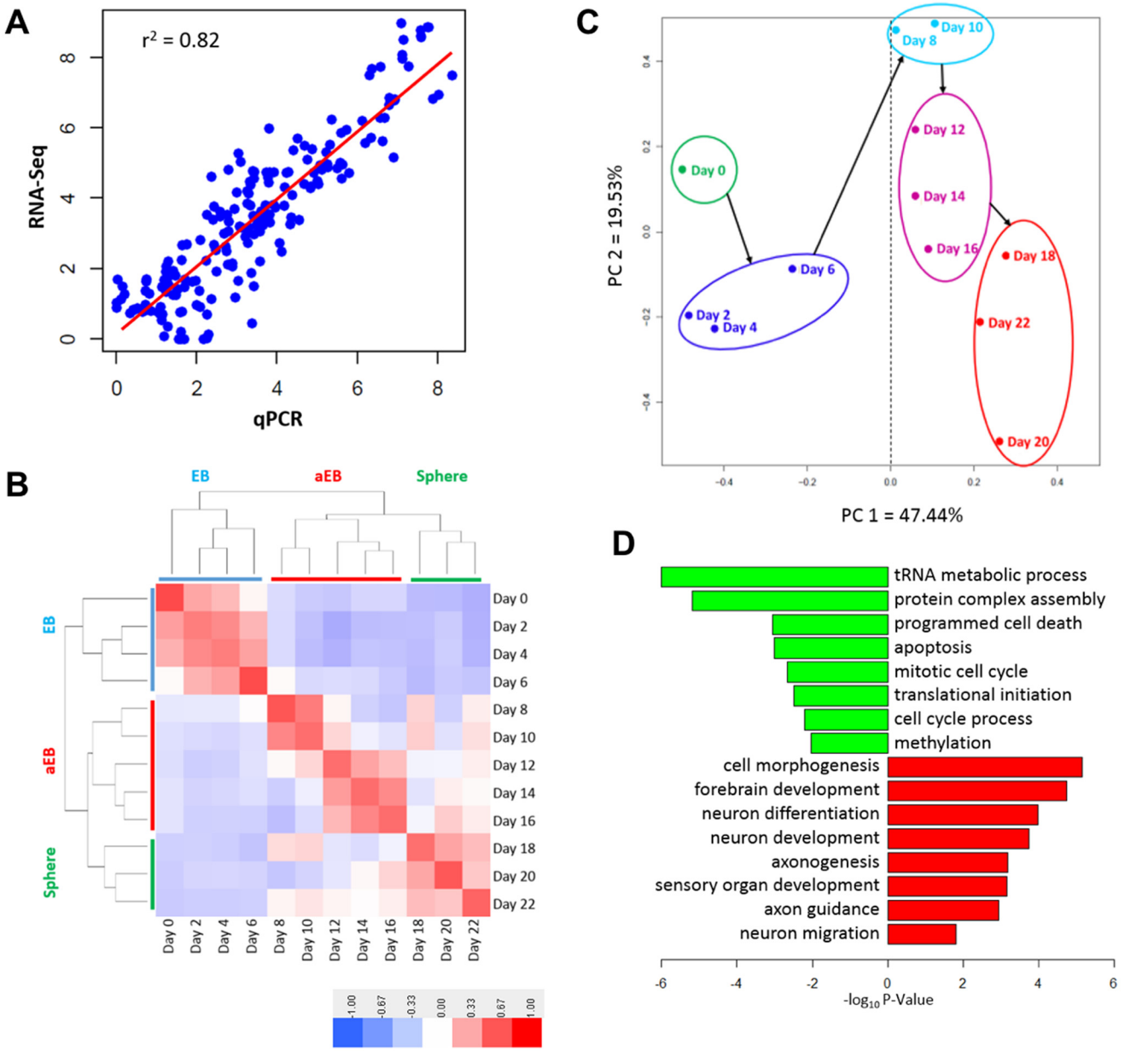
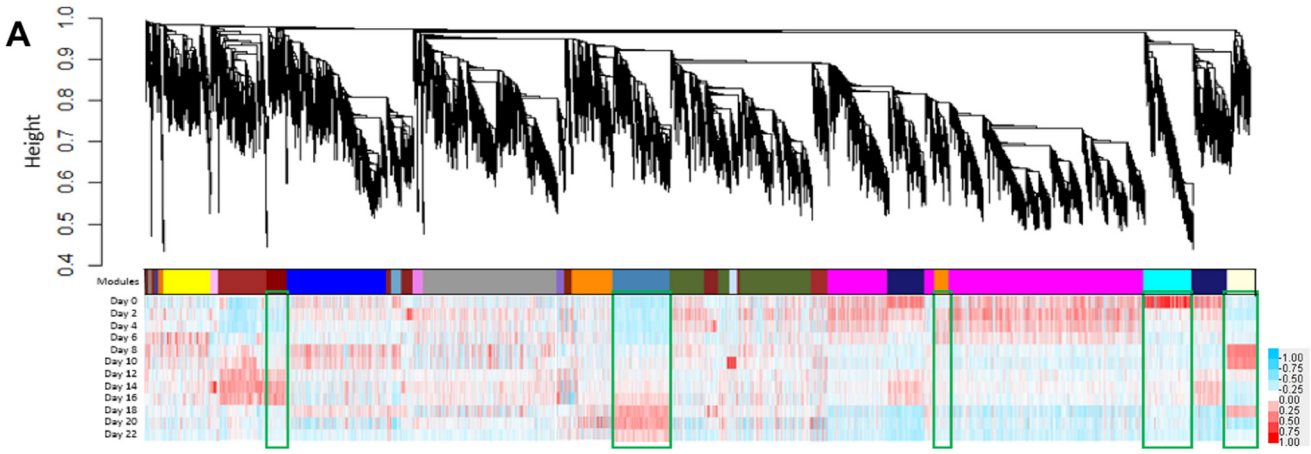


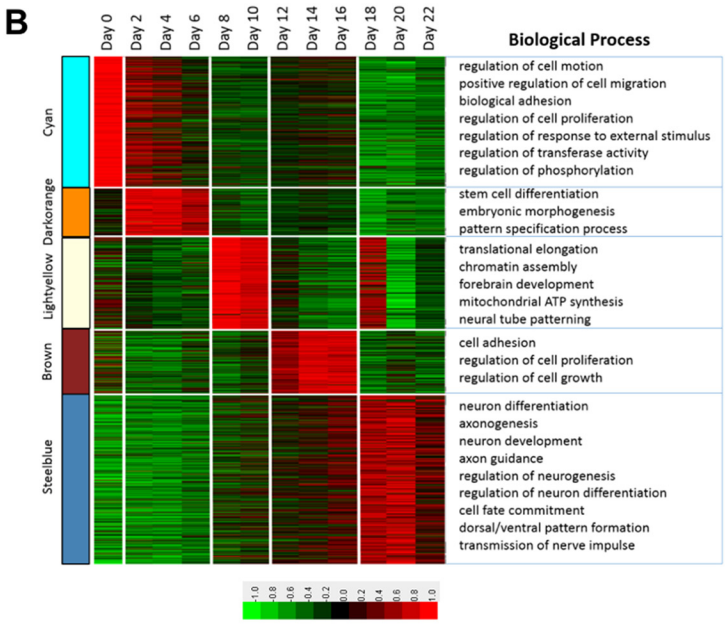
Figure 2. Time-series transcriptome analysis of hESC *in vitro* neural differentiation. *A,* correlation between gene expression levels measured by RNA-Seq and RT-qPCR for marker genes at the different time points. *B,* unsupervised hierarchical clustering analysis of all 12 population samples. *C,* principal component analysis of transcriptomes of all 12 population samples. *D,* GO enrichment analysis for PC1-positive and -negative genes.

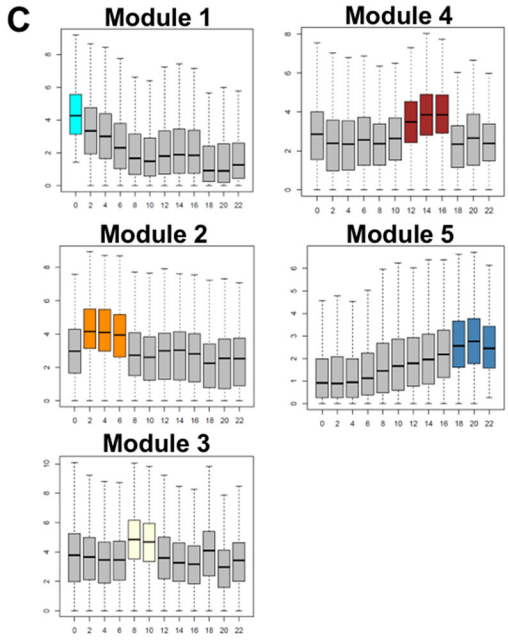
The expression of genes in Module 1 was down-regulated, whereas the expression of genes in Module 5 was up-regulated gradually along the differentiation of hESCs (Fig. 3, *A*–*C*).

To characterize five modules individually, functional enrichment analysis was conducted for genes in each module. Module 1 (day 0) was enriched for pluripotency related terms, *e.g.* regulation of cell motion, migration, and biological adhesion. Module 2 (day 2/4/6) was associated with stem cell differentiation and embryonic morphogenesis. The terms of chromatin assembly, forebrain development, and neural tube patterning were identified in Module 3 (day 8/10). Cell proliferation and cell growth were enriched in Module 4 (day 12/14/16). Module 5 (day 18/20/22) was associated with neuron differentiation, axonogenesis, and neuron development (Fig. 3*B*). To chart the activity of signaling pathways in the different modules, we examined the enrichment of target genes of important signaling pathways by using RNA-seq or microarray data from published perturbation experiments (29), and we found many signaling pathways were involved during hESC neural differentiation (Fig. 3*D*); some pathways are activated (*red*) and some are inhibited (*green*) in different modules, but there was barely any detectable signaling activity in Module 3, which is associated with day 8/10 (Fig. 3*D*). The enrichment of neural related genes and the lack of extrinsic signaling genes in Module 3 support that day 8/10 is a critical period for the fate transition from pluripotency to neural lineage.

Based on the signature gene profiles in the defined modules, the early neural differentiation of hESC can be classified into five stages, including pluripotency stage with Module 1 at day 0, differentiation initiation stage with Module 2 at day 2/4/6, neu-



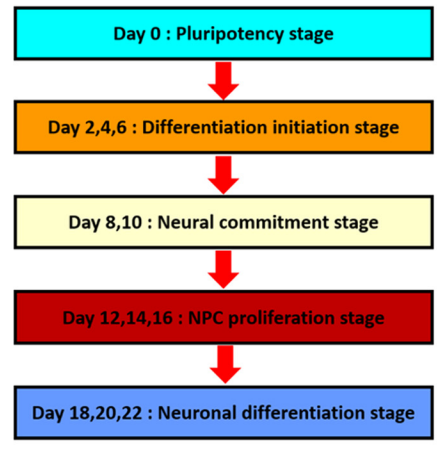




D

PATHWAY	Module1	Module2	Module3	Module4	Module5
TGFB_A	11.98	0.21	1.33	19.79	2.13
TGFB_I	6.18	3.80	2.32	2.11	3.26
BMP_A	34.98	3.73	1.16	12.07	0.23
BMP_I	2.46	0.67	0.23	0.23	0.41
HEDGEHOG_A	0.79	1.83	1.32	3.78	14.72
HEDGEHOG_I	5.05	3.71	0.04	1.83	1.64
FGF_A	0.21	0.97	0.00	1.33	3.29
FGF_I	3.58	0.90	1.42	2.91	1.33
HIPPO_A	23.11	0.20	0.04	6.19	0.92
HIPPO_I	2.54	3.58	0.21	0.25	0.36
NODAL_A	3.73	0.49	1.83	2.65	1.29
NODAL_I	2.83	3.34	0.02	3.41	1.33
NOTCH_A	2.98	0.11	0.23	4.64	1.01
NOTCH_I	0.03	1.13	0.54	2.90	5.05
WNT_A	6.00	0.49	0.23	0.00	0.29
WNT_I	1.76	0.40	0.04	2.50	0.04

Activate (A) Inhibit (I)



ral commitment stage with Module 3 at day 8/10, NPC proliferation stage with Module 4 at day12/14/16, and neuronal differentiation stage with Module 5 at day18/20/22 (Fig. 3*E*). Taken together, we reveal that the hESC neurodevelopmental trajectory contains five stages, characterized by unique module genes, and day 8/10 is associated with neural commitment.

Comparative analysis validates Module 3 associated day 8/10 stage as a critical period for the neural commitment

In mouse early embryo development, the neural fate commitment occurs during gastrulation, and a portion of anterior ectoderm is specified to adopt the neural fate (28, 29). One of our previous studies documented the spatial transcriptome of the epiblast at the mid-gastrulation (E7.0) (29). Pearson correlation coefficient (PCC) assays with RNA-Seq data of all the embryo regions and five hESC differentiation stages showed that Module 3 was specifically enriched with the prospective ectoderm in the anterior embryo at E7.0, where the neural epithelium originates later (Fig. 4*A*). In contrast, Module 2 was widely associated with the whole epiblast, indicating the status of the pluripotency (Fig. 4*A*). In mouse ESC neural differentiation, the neural fate commitment begins when NPC-specific gene *Sox1* starts to express (30). By comparing with the microarray data of *in vitro* mouse ESC neural differentiation, we noticed that day 8 of Module 3 was highly correlated with the day 3 $Sox1^+$ cells, indicating the initiation of mouse neural fate (Fig. 4*B*). Meanwhile, Module 2 was associated with cells in pluripotent status (Fig. 4*B*). All the findings above demonstrated that neural fate commitment occurs at day 8/10 stage during hESC differentiation.

The CORTECON system is generated by analyzing RNA-Seq data with nine samples from day 0 to day 77 of the differentiating hESCs (20). PCC assays revealed that the CORTECON system could not reveal the conversion from the pluripotent epiblast to the neural lineage, which progresses quickly during the early development of the human nervous system (Fig. 4*C*). Transcriptome profiling data, generated by microarray from various laser-microdissected human fetal brain regions at different stages (31), are available in the BrainSpan Atlas of the Developing Human Brain (http://brainspan.org/).[4] The correlation analysis demonstrated that our hESC neural differentiation system resembles the development of early human brain at gestational week 8/9 (Fig. 4*D*).

Next, we asked whether there were novel markers that could be used to distinguish each stage. We performed Guilt-by-Association analysis (24, 32) to identify putative sub-stage–specific markers (Fig. 4*E*). As shown in supplemental Fig. S3, the expression of putative markers in each stage was activated sequentially following the differentiation of hESCs. Then, RT-qPCR assays were conducted to confirm the expression profiles of selected genes in each module (supplemental Fig. S4*A*). *GAL*, *GNA14*, *CALB1*, and *TUBA4A* were novel marker genes in Module 1 for pluripotency; *SALL3*, *PCDH1,* and *IRX2* in Module 2 for differentiation initiation; *FGF8* and *PTN* in Module 3 for neural fate commitment; *ENPP2*, *SEMA5A,* and *NEDD9* in Module 4 for NPC proliferation; and *STMN2* and *EMX2* in Module 5 for neuronal differentiation (supplemental Fig. S4, *A* and *B*). The expression patterns of those novel marker genes with known stage-specific genes, *e.g. OCT4*, *OCT6*, *PAX6*, *SOX1,* and *NCAM,* during the hESC neural differentiation are summarized in supplemental Fig. S4*B*.

Overall, systematic comparison between our RNA-Seq data and other transcriptome profiling data suggests that day 8/10 stage is a critical window for the fate transition from the pluripotency to the neural lineage.

Regulation network of transcription factor genes during hESC neural differentiation

Given that intrinsic regulators, especially TFs, play essential roles in the neural commitment (20, 33, 34), we sought to identify key TF genes in each of the five modules. By overlapping the stage-specific genes with TF database (35) (www.transcriptionfactor.org),[4] distinct TF genes were enriched in individual modules (Fig. 5*A*). The gene ontology analysis showed that the stage transition could possibly be driven by these TFs during the hESC neural differentiation. For instance, TFs in Module 1 were enriched for the regulation of transcription, gene expression, cellular biosynthetic process, and cell proliferation (supplemental Fig. S5*A*). TFs in Module 2 were related to embryonic morphogenesis, tube development, tissue morphogenesis, and ectoderm development (supplemental Fig. S5*B*). TFs in Module 3 were associated with forebrain development, cell fate commitment, neural tube development, and midbrain–hindbrain development (supplemental Fig. S5*C*). TFs in Modules 4 and 5 were involved in cell proliferation and neuronal differentiation (supplemental Fig. S5, *D* and *E*).

To unravel the inter-group connection of TFs in each module, we analyzed the Connection Specificity Index (CSI) (36) of the five TF groups, and generated a TF co-expression network (CSI >0.9). The positive correlations between the different TFs are marked in *red*, and the negative correlations are marked in *green* (Fig. 5*B*). TFs in each module showed positive interactions with each other (Fig. 5*B*), indicating that TFs within the module may form a synergistic regulatory circuitry and act in a combinatory manner. TFs in Modules 1 and 2 mainly showed positive correlation, whereas TFs between Modules 1 and 5 displayed negative correlation. Interestingly, TFs in Module 3 that presented the critical period of the neural fate transition displayed unique correlations with other TF groups (Fig. 5*B*). They were mainly negatively correlated with TFs in Modules 1 and 2, inhibiting earlier processes; they were positively corre-

[4] Please note that the JBC is not responsible for the long-term archiving and maintenance of this site or any other third party hosted site.

Figure 3. WGCNA revealed gene-network modules enriched in hESC neural differentiation. *A,* hierarchical cluster tree showing co-expression modules identified using WGCNA. Modules correspond to branches and are labeled by colors. *B,* heat map showing the relative expression of genes in five representative stage-specific modules. Top gene ontology terms are shown on the *right side*. *C, box plots* showing the distribution of module expression (median FPKM of all genes within a given module) for different substages. *D,* analysis of signaling pathway involved in each module. *Red* and *green* indicate activated enrichment and inhibitory enrichment, respectively (*I,* inhibit; *A,* activate). The value in each cell is the log10(FDR) and the *red/green border* indicates significant enrichments (FDR <0.01). *E,* schematic drawing of the five sequential substages of hESCs neural differentiation.

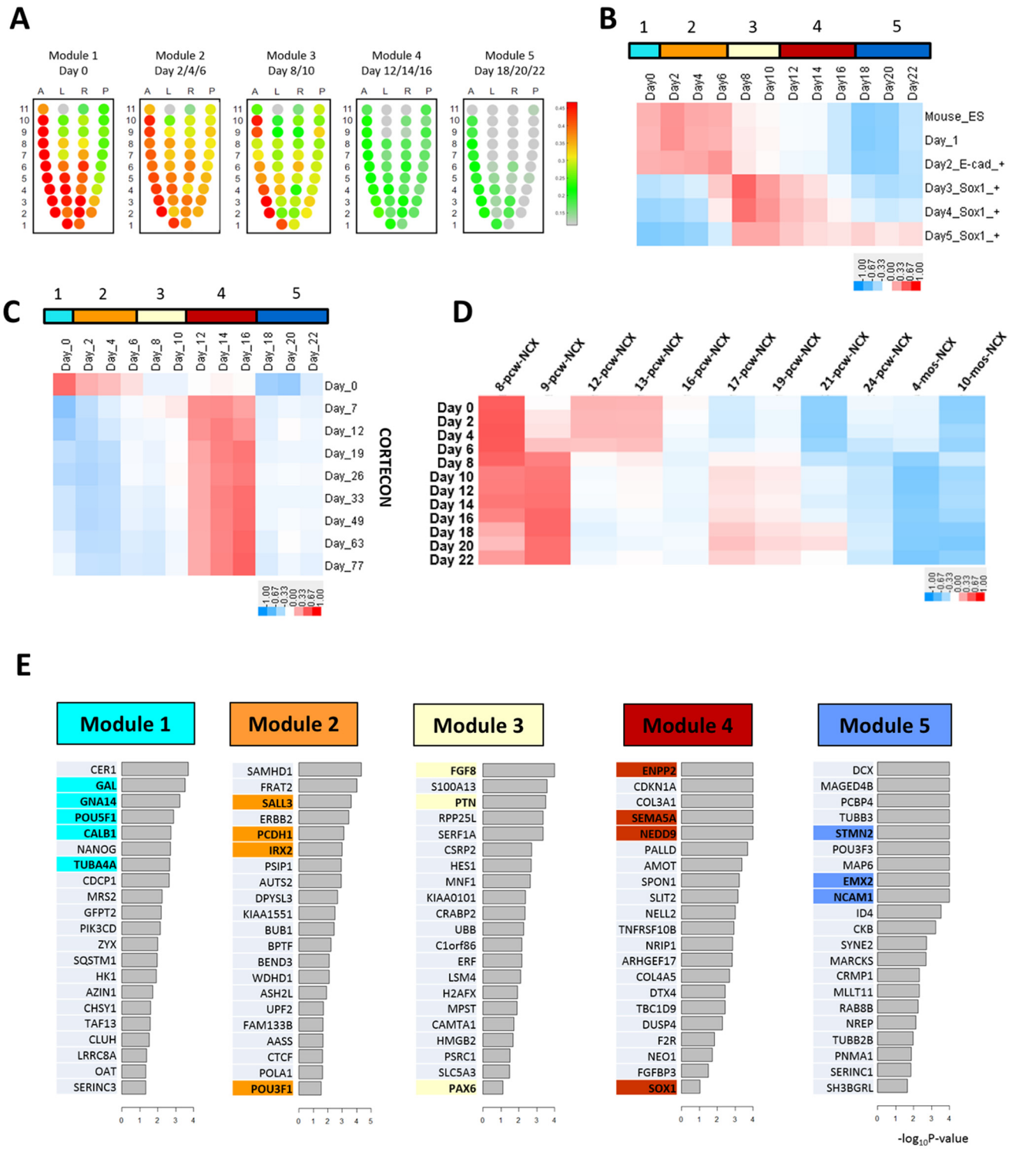


Figure 4. Comparative analysis validates Module 3 associated day 8/10 stage as a critical period for the neural commitment. *A,* comparative analysis of the hESC neural differentiation RNA-Seq data with laser-microdissected E7.0 mouse embryo regions. *B,* comparison analysis of the hESC neural differentiation RNA-Seq data with *in vitro* mouse ESC neural differentiation microarray data. *C,* comparison analysis of the hESC neural differentiation RNA-Seq data with CORTECON neural differentiation system. *D,* comparative analysis of the hESC neural differentiation RNA-Seq data with BrainSpan Atlas of the Developing Human Brain of transcriptome profiles from laser-microdissected human fetal brain regions. *E, bar graphs* showing the top 20 putative marker genes for each module. A few candidate markers, which are *bold* and *colored*, were validated by RT-qPCR.

lated with TFs in Modules 4 and 5, promoting later events, suggesting that they might behave like a gatekeeper to mediate the lineage transition from the pluripotency to the neural fate during the hESC differentiation (Fig. 5*B*).

TFs with the highest degree of connectivity in all modules are termed hub genes (31), which are expected to play imperative roles during the differentiating process. Hub genes of each module were identified and listed in Fig. 5*C*. *PAX6*, *SIX3*, *SIX6*,

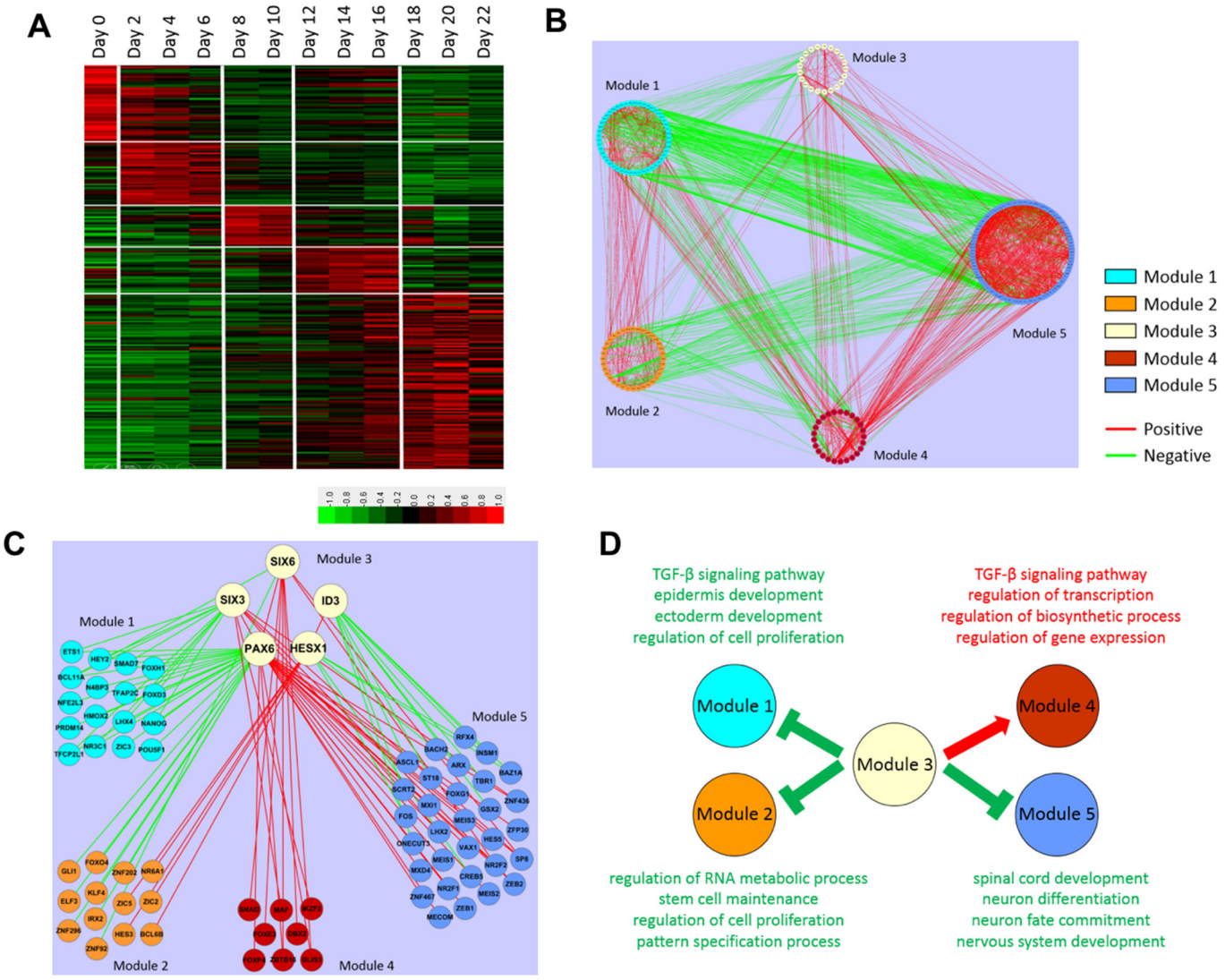


Figure 5. Regulation network analysis of transcription factors in each stage. *A,* heatmap showing the expression pattern of transcription factors in each substage of hESC neural differentiation. *B,* CSI network of all the TFs. TFs within the same module are grouped together and colored. Gene expression PCC-derived CSIs are calculated based on the RNA-Seq expression values. *C,* regulation network of Top 5 hub genes in Module 3. *D,* model for the role of Module 3 (neural commitment stage) in the hESC neural differentiation.

HESX1, and *ID3* were the top hub TFs with the highest CSI in Module 3 (Fig. 5*C*). In mice, both *Pax6* and *Six3* genes are essential for the appropriate forebrain patterning (37–40). As expected, *PAX6* and *SIX3* centering in the hub generated the most connections with other modules (Fig. 5*C*). Both *PAX6* and *SIX3* were negatively correlated with hub TF genes in Module 1 and/or 2 but positively associated with hub TF genes in Modules 4 and 5 (Fig. 5*C*). *SIX3* and *SIX6* displayed similar regulation patterns with *PAX6* (Fig. 5*C*). However, *HESX1* and *ID3* were negatively correlated with the hub genes in Module 5, and *HESX1* positively correlated with the hub genes in Module 2. Thus all the data above support a notion that Module 3, which is associated with the neural fate commitment, inhibits Modules 1, 2, and 5, but it promotes Module 4 (Fig. 5*D*). The hub TF analysis also indicated that the cooperation of key TFs not only governs the progress in each sub-stage but also mediate the transition between temporal stage transition, which generates a dynamic and hierarchical TF regulatory network during hESC neural differentiation.

SIX3 and HESX1 play imperative roles in hESC neural differentiation

Among five hub genes of Module 3, the functions of *PAX6* in human neurodevelopment have been well studied (41, 42). To validate the function of other hub TF genes in Module 3, the CRISPR/Cas9 gene edition system was used to target *SIX3*, *SIX6*, *HESX1*, and *ID3* in hESCs. Unfortunately, we failed to generate a stable hESC colony with either *SIX6* or *ID3* gene deletion (data not shown), indicating that *SIX6* and *ID3* genes might be essential for the maintenance of hESCs. The sgRNAs with complementary sequence to the coding sequence of the *SIX3* or *HESX1* gene were used to achieve gene edition (supplemental Table S2). Two individual colonies were screened for the editing of each gene. The sequencing data revealed that *SIX3*-sgRNA #1 (*SIX3* #1 for short) and *SIX3*-sgRNA #2 (*SIX3* #2 for short) had a 1-bp deletion at different loci of the *SIX3* genome, and *HESX1*-sg RNA #1 (*HESX1* no. 1 for short) and *HESX1*-sgRNA #2 (*HESX1* #2 for short) had the same 12-bp

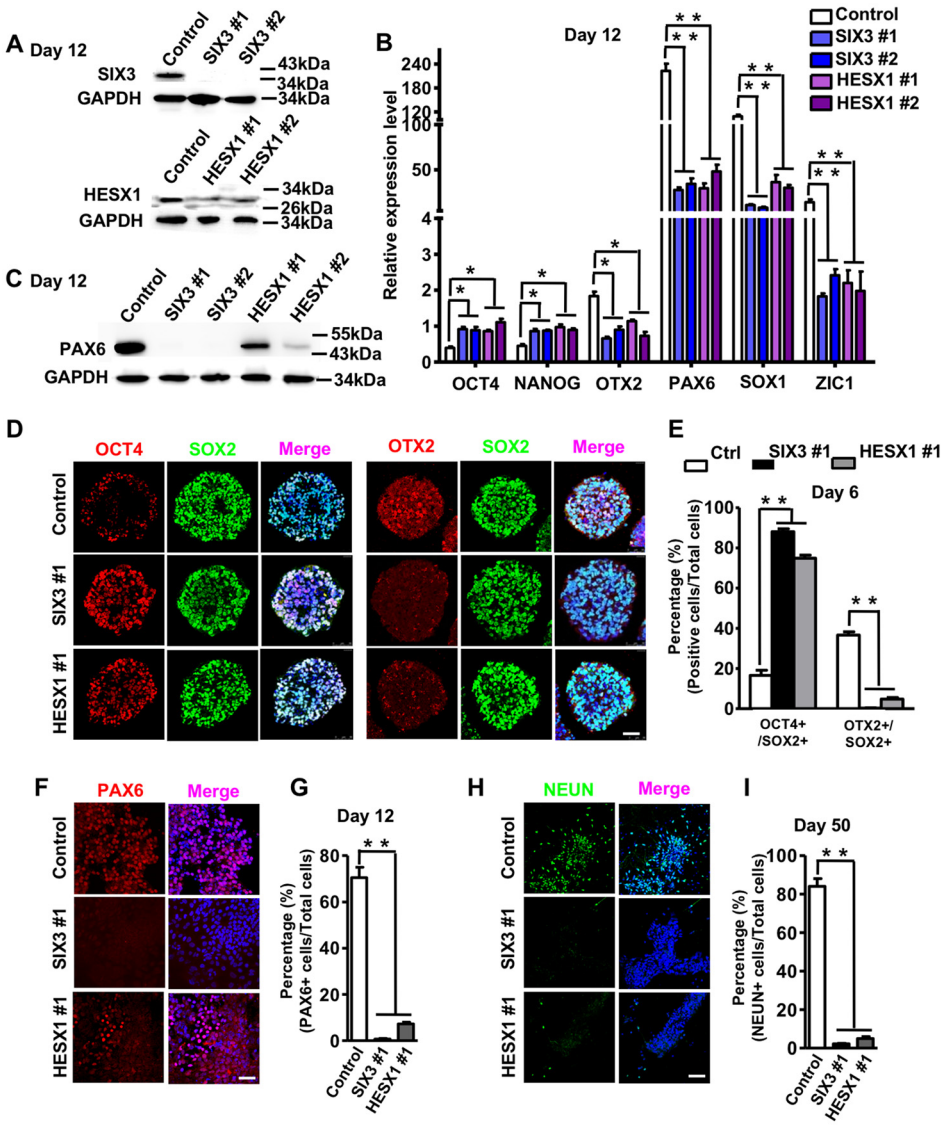


Figure 6. Generation of SIX3 and HESX1 knock-down hESC lines using CRISPR/Cas9 system. *A,* Western blot for CRISPR/Cas9 efficiency validation of *SIX3* and *HESX1* knockdown at day 12 differentiation samples. SIX3 is about 37 kDa; HESX1 is about 47 kDa, and GAPDH is 34 kDa. *B,* RT-qPCR analysis for the gene expression at day 12 in the *SIX3* knock-out and *HESX1* knockdown cell lines. *C,* Western blot for CRISPR/Cas9 efficiency validation of *PAX6* knockdown at day 12 differentiation samples. PAX6 is about 50 kDa. *D,* double immunocytochemistry analysis of *OCT4* and *SOX2*, *OTX2* and *SOX2* respectively, on day 6. *Scale bars,* 75 μm. *E,* quantification of the immunocytochemistry in *D*. *F,* immunocytochemistry analysis of *PAX6* on day 12. *Scale bars,* 75 μm. *G,* quantification of the immunocytochemistry in *F*. *H,* immunocytochemistry analysis of *NEUN* on day 50. *Scale bars,* 75 μm. *I,* quantification of the immunocytochemistry in *H*. To quantify the differentiation efficiency, three to five fields were randomly selected. $n = 3$ independent experiments; two-tailed *t* test. All data are presented as the mean ± S.D. *, $p < 0.05$; **, $p < 0.01$. Related to *E, G,* and *I*.

deletion (data not shown). The Western blot assays confirmed that compared with the mock control at differentiation day 12, the expression of the SIX3 protein was barely detected in both colonies, and the expression of the HESX1 protein was significantly decreased in the two edited colonies (Fig. 6*A*). The two gene-edited colonies of either *SIX3* or *HESX1* gene could be passaged normally; in addition, the expression of *OCT4*, *SOX2*, and *NANOG* genes was comparable between those colonies and the control at both RNA and protein levels (supplemental Fig. S6, *A* and *B*), suggesting that these genes do not affect the pluripotency maintenance.

RT-qPCR assays revealed that at differentiation day 12, the expression of pluripotent genes *OCT4* and *NANOG* was increased, but the neural epithelium marker genes *OTX2*, *PAX6*, *SOX1*, and *ZIC1* was significantly reduced (Fig. 6*B*). Consistently, Western blot assay confirmed that the expression of the PAX6 protein in the gene-edited colonies of either gene was also reduced (Fig. 6*C*).

Next, we asked what were the consequences of loss-of-function of the *SIX3* or *HESX1* gene during hESC neural differentiation. Immunocytochemical assays were performed to detect the expression of OCT4, SOX2, and OTX2 at day 6 EBs, and of PAX6 at day 12 attached EBs, and of NEUN at day 50 neurons in the control and *SIX3* KO or *HESX1* KO colony. In day 6 EBs, compared with the control, there were more $OCT4^+/SOX2^+$ cells in either the *SIX3* KO or *HESX1* KO colony; but much fewer $OTX2^+/SOX2^+$ cells (Fig. 6, *D* and *E*), suggesting that the early neural differentiation process was inhibit by loss-of-function of *SIX3* or *HESX1*. In day 12 attached EBs, compared with the control, there were much fewer $PAX6^+$ cells in either *SIX3*

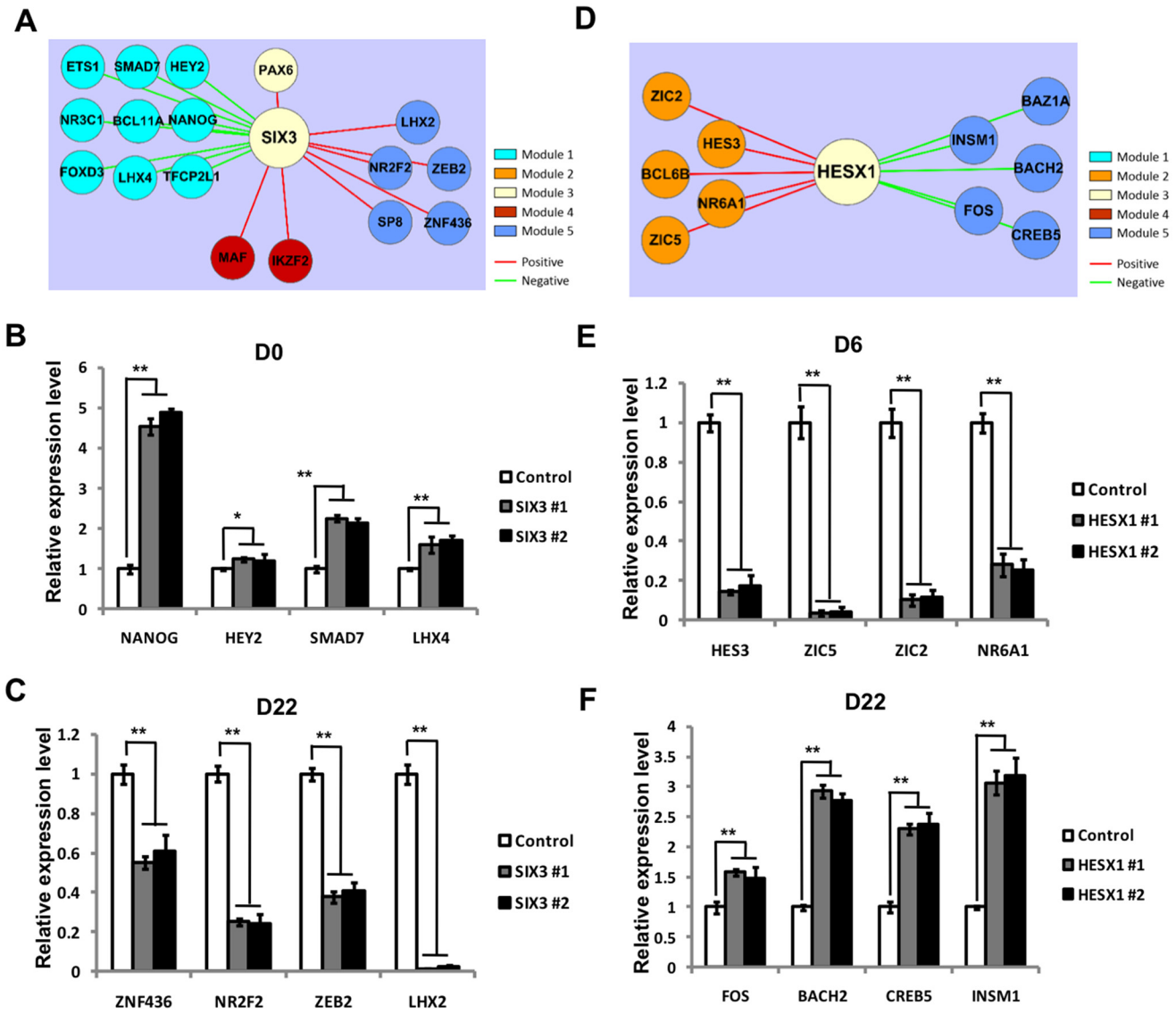


Figure 7. qPCR analysis of downstream target gene of *SIX3* and *HESX1* in CRISPR/Cas9 knock-out cell lines. *A,* regulation network of hub TF *SIX3*. *B,* relative expression level of *SIX3* downstream target genes in KO cell lines at day 0. *C,* relative expression level of *SIX3* downstream target genes in KO cell lines at day 22. *D,* regulation network of hub TF *HESX1*. *E,* relative expression level of *HESX1* downstream target genes in KD cell lines at day 6. *F,* relative expression level of *HESX1* downstream target genes in KD cell lines at day 22. For q-PCR, a minimum of three biological replicates from two separate experiments were examined.

KO or *HESX1* KO colony (Fig. 6, *F* and *G*). In day 50 neurons, there were many NEUN$^+$ cells in the control; however, in either the *SIX3* KO or *HESX1* KO colony, very few NEUN$^+$ cells were detected (Fig. 6, *H* and *I*), suggesting that neuronal differentiation was blocked by *SIX3* or *HESX1* KO. Similar results were obtained in another clone of *SIX3* and *HESX1* deletion cells (supplemental Fig. S7). Taken together, these data suggest that the neural differentiation of hESCs is compromised in the loss-of-function of either the *SIX3* or *HESX1* gene.

SIX3 and HESX1 promote neural differentiation by regulating downstream TF networks

According to the correlation analysis of hub genes among different modules (Fig. 5*C*), the potential downstream target genes of either *SIX3* or *HESX1* were identified (Fig. 7, *A* and *D*). If *SIX3* or *HESX1* promotes hESC neural differentiation through its downstream target genes, the loss-of-function of *SIX3* or *HESX1* gene should cause the up-regulation of its negatively correlated genes but the down-regulation of its positively correlated genes. Indeed, the expression of *SIX3* negatively correlated genes *NANOG*, *SMAD7*, and *LHX4* in Module 1 was significantly increased in both *SIX3* KO colonies at day 0 (Fig. 7*B*); in contrast, the expression of positively correlated genes such as *NR2F2*, *ZEB2*, and *LHX2* was reduced (Fig. 7*C*). In both *HESX1* KO colonies, similar expressing changes were observed for the downstream target genes of *HESX1*, including *HES3*, *ZIC5*, *ZIC2*, and *NR6A1* in Module 2 (Fig. 7*E*), as well as *FOS*, *BACH2*, *CREB5*, and *INSM1* in Module 5 (Fig. 7*F*). These results suggest that *SIX3* and *HESX1* genes may intrinsically promote neural differentiation by regulating its downstream TF networks.

Discussion

In this study, we establish an *in vitro* hESC neural differentiation system, which mimics the early human brain develop-

ment. Twelve RNA samples were collected from differentiating hESCs every other day between day 0 and day 22 for RNA-Seq. Transcriptome analysis reveals that five distinct co-expression gene modules can classify the hESCs neural differentiation into five sub-stages. The gene expression profiling and CSI analysis of transcription factor correlation suggest that Module 3 associated day 8/10 stage is a critical period for the neural fate commitment. *PAX6*, *SIX3*, *SIX6*, *HESX1*, and *ID3* are identified as key hub TF genes of Module 3. CRISPR/Cas9–mediated functional assays reveal that the neural differentiation of hESCs is impaired with the loss-of-function of either *SIX3* or *HESX1* gene.

Many studies with mouse systems have contributed significantly to the understanding of the development and the function of mammalian brain (43–45). Nevertheless, there are notable differences between the mouse and human brains, such as a 1,000-fold difference in size, more complicated human cortical layers, and advanced emotion- and stress-related functions of the human brain. Especially, mouse models of neural diseases cannot fully recapitulate the characteristics of human disorders (46). Neural differentiation of hESCs has been demonstrated to be a powerful system to study human brain development (47–51). In this study, we adapted and established an *in vitro* hESC neural differentiation system (Fig. 1). Similar to previous reports (14, 15), RT-qPCR and immunofluorescent assays demonstrate that our hESC neural differentiation system is reliable and reproducible to resemble the early human cortical development *in vivo* (Fig. 1 and supplemental Fig. S1). The early neural development of hESCs progresses very quick (13, 15, 23). The insufficient time points analyzed by RNA-Seq or the low resolution of the microarray technique in some previous studies limits the precision analysis of early neurodevelopmental events (20, 22). Our transcriptome analysis, based on RNA-Seq data of 12 samples, reveals that during the early neural differentiation of hESCs, there are five temporal stages as follows: pluripotency (day 0); differentiation initiation (day 2/4/6); neural commitment (day 8/10); NPC proliferation (day 12/14/16); and neuronal differentiation stage (day 18/20/22) (Figs. 2 and 3). PCC assays with our RNA-Seq data and hESC CORTECON show that day 8/10 stage may mediate the fate switch from the pluripotent epiblast to the neural epithelium (Fig. 4). All the evidence above suggests that our hESC neural differentiation system is suitable to investigate the molecular and cellular mechanisms of early human brain development such as neural commitment.

In the past century, neural commitment has been well studied in amphibian and rodents (8, 10). The inhibition of BMP signaling seems revolutionarily conserved during the neural fate initiation in different organisms (9, 13, 52). Inhibition of the TGFβ pathway has now been demonstrated to be sufficient to directly induce neural fate in mammalian embryos as well as pluripotent mouse and human embryonic stem cells (53), and consistently, our data also showed that the TGFβ signaling has been inhibited during hESC neural differentiation processing (Fig. 3*D*). However, the involvement of different extrinsic signaling pathways and intrinsic factors, including transcriptional and epigenetic factors, suggests that the regulatory machinery of neural determination is not that simple. So far, how the neural fate is initiated in humans has not been clearly elucidated. In our study, the unique expression profiles of genes in PCA and WGCNA analysis suggest that among the five temporal substage, the Module 3 associated day 8/10 stage is a critical period for the fate transition (Figs. 2 and 3 and supplemental Fig. S2). Functional enrichment analysis reveals that Module 3 is associated with forebrain development (Fig. 3); in addition, TFs in Module 3 are related to cell fate commitment and neural tube development (Fig. 5 and supplemental Fig. S5*C*). Moreover, systematic comparison between our RNA-Seq data and other transcriptome profiling data such as mouse Geo-Seq (29), mESC microarray (22), hESC CORTECON (20), and human fetal brain microarray (31) suggests that day 8/10 stage is essential for the neural determination (Fig. 4). Furthermore, no active extrinsic signals are present in Module 3 of day 8/10 (Fig. 3), which supports the notion that the neural commitment is ensured by the intrinsic program. Finally, TFs in Module 3, as well as hub TFs, including *PAX6*, *SIX3*, *SIX6*, *HESX1*, and *ID3,* are identified (Fig. 5 and supplemental Table S1). The CSI assays show that TFs in Module 3 are negatively correlated with TFs in Modules 1 and 2 but are positively associated with TFs in Modules 4 and 5 (Fig. 5). As expected, the loss-of-function of either *SIX3 or HESX1* gene leads to the compromised neural differentiation of hESCs (Fig. 6 and supplemental Fig. S7), which will be discussed in detail later. Thus, all our findings above demonstrate that day 8/10 stage is a critical window for the fate transition from the pluripotency to the neural lineage.

The intrinsic regulation is imperative for the neural induction. Nonetheless, how the key intrinsic factors, especially transcriptional factors, participate in the regulation of neural commitment is still largely unclear. The determination of day 8/10 stage as a neural fate conversion period provides an opportunity to explore the mechanism of human neural induction. *PAX6*, *SIX3*, *SIX6*, *HESX1*, and *ID3* are hub TF genes in Module 3 related to day 8/10 (Fig. 5). It has been known that *PAX6* is not only a marker of human neural stem cells but is also a human neuroectoderm cell fate determinant (41, 54, 55). In humans, various mutations of the *SIX3* gene, mapped to 2q21, are associated with holoprosencephaly (56–60) and schizophrenia (61). Different mutations of the *HESX1* gene, mapped to 3p14.3, are related to septooptic dysplasia (62–65) and pituitary deficiency (63, 66–68). In addition, the expression of *Six3* and *Hesx1* genes is highly enriched in the prospective E7.0 mouse anterior ectoderm (29), where the prospective neural epithelium originates, suggesting that *SIX3* and *HESX1* may play a crucial role in human neural commitment as well. Indeed, the early neural differentiation of hESCs is repressed with either *SIX3* or *HESX1* deficiency, achieved by the CRISPR/Cas9 gene edition system (Fig. 6 and supplemental Fig. S7). Several possible downstream target genes of either *SIX3* or *HESX1* are identified by the correlation analysis. As expected, in the *SIX3*-deficient cells, the expression of the negatively associated genes such as *NANOG*, *SMAD7,* and *LHX4* is increased; in contrast, the expression of the positive correlated genes, including *NR2F2*, *ZEB2,* and *LHX2,* is decreased (Fig. 7, *A–C*). Similar results are obtained for the *HESX1*-regulated genes in the *HESX1*-deficient cells (Fig. 7, *D–F*). Thus, our findings provide evidence that *SIX3* and *HESX1* are novel key intrinsic factors,

which ensure human neural commitment by regulating the distinct transcriptional network. Our system can be used not only to identify the pivotal determinants in temporal events during human brain development, but also to unravel their mechanisms to program the appropriate neurodevelopment and to prevent the neural diseases.

Our *in vitro* hESC neural differentiation system recapitulates early human cortical development *in vivo*; especially the day 8/10 stage resembles the neural commitment in human brain. Many neural diseases such as autism spectrum disorders, schizophrenia, depression, epilepsy, and mental retardation are caused by the abnormal early neurodevelopment. Thus, our findings will benefit not only the understanding of the molecular mechanism of human brain development, but also the understanding of the etiology and the therapy of human mental diseases.

Experimental procedures

Human ESC culture

Human ESC lines H9 and H1 (passage 30–45) were maintained and passaged every 6–7 days on a feeder layer of irradiated embryonic mouse fibroblasts as described previously (12, 69). The standard medium was supplemented with 4 ng/ml fibroblast growth factor (R&D Systems Inc.), and the differentiated colonies were physically removed before passaging.

Human ES cell neural differentiation

The procedure for hES cell differentiation is schematically summarized in Fig. 1*A*. Briefly, human ES cell colonies were detached from the feeder layer and floated in the ES cell growth medium without basic FGF. EB were formed after 4 days' suspension in the medium, and subsequently floated for 2 days in neural induction medium consisting of F-12/DMEM, N2 supplement, and non-essential amino acids (NEAA). The ES cell aggregates (EBs) were then adhered to a substrate in a neural induction medium. By about 10 days after ES cell differentiation, cells in the center of each colony differentiated into neuroectodermal cells, displaying small columnar morphology followed by organization of the columnar cells into neural tube-like rosettes after an additional 6 days. These neural epithelial cells were isolated from surrounding non-neural cells through differential response to dispase treatment. These aggregates of neural epithelial cells are floating and cultured in the same neural induction medium for another 6 days to form the neural sphere. Finally, these the neural progenitor spheres were digested with accutase (Innovation Cell Technology) into single cells and plated on PDL-coated dishes in the N2B27 medium for maturation (12).

Immunocytochemistry and cellular quantification

Immunocytochemistry staining on dish cultures was performed as described previously (15). The primary antibodies used in this study are listed in supplemental Table S3.The cells were fixed with 4% paraformaldehyde in PBS (pH 7.3) for 2 h. Primary antibodies used in this research included the following: OCT4 (Santa Cruz Biotechnology, catalog no. sc-5279); SOX2 (Abcam, catalog no. ab59776); NANOG (CST, catalog no. 8822); TRA-1–81 (Chemicon & Millipore, catalog no. 90233) and SSEA3/4 (Chemicon & Millipore, catalog no. 90231); NESTIN (Abcam, catalog no. AB6142); OTX2 (catalog no. AF1979) and SOX1 (R&D Systems, catalog no. AF3369); MAP2 (Sigma, catalog no. M4403); NEUN (catalog no. ABN78), VGLUT1/2 (catalog no. MAB5502), GAD 67(catalog no. AB1511), and TH (Chemicon & Millipore, catalog no. 657012); VAChT (Synaptic Systems, catalog no. 139103); and TBR1 (Abcam, catalog no. AB31940) (supplemental Table S3). Images were taken with Nikon Eclipse FN1 confocal laser-scanning microscope. To quantify the differentiation efficiency, three to five fields were randomly selected, and the neuronal cells and total cells (DAPI-stained) were counted using ImageJ software. All the experiments were performed in triplicate. A two-tailed Student's *t* test was used for statistical analysis.

Quantitative reverse-transcription PCR (qRT-PCR)

Total RNA was extracted using TRIzol reagent (Invitrogen), and cDNA was reverse-transcribed using the SuperScript III First-strand cDNA synthesis kit (Invitrogen). Quantitative RT-PCR was performed using the MyiQ real-time PCR detection system (Bio-Rad), as described (14). A minimum of three biological replicates from two separate experiments were examined.

Single-cell qPCR

The single-cell qPCR follows the protocol as published previously (70). Briefly, the single cell is collected into lysis buffer for the first cDNA synthesis, and a modified Smart2-seq protocol is performed to amplify the cDNA. cDNAs that pass quality control are then used for qPCR. The procedure has been optimized for each step: tissue collection; cell lysis; RNA isolation; and single cell-based PCR amplification. Quantitative RT-PCR was performed using the MyiQ real-time PCR detection system (Bio-Rad), as described (14). A minimum of three biological replicates from two separate experiments were examined.

CRISPR/Cas9 gene edition

The single-cell qPCR follows the protocol as published previously (71). First, *pX330* plasmid, including the T7 promoter, was linearized by NotI. Linearized templates were purified and transcribed *in vitro* with mMESSAGE mMACHINE T7 ULTRA kit (Life Technologies, Inc.). sgRNAs with the T7 promoter were amplified by PCR and transcribed *in vitro* using MEGAshortscript T7 kit (Life Technologies, Inc.). After transcription, the Cas9 sgRNAs were purified with MEGAclear kit (Life Technologies, Inc.) according to the manufacturer's instructions. The sequences for preparation of template for *in vitro* transcription of sgRNA are shown in supplemental Table S2.

Western blotting

Protein samples were collected on day 12 of differentiation. The harvested cells were lysed with 1× loading buffer and heated for 10 min at 100 °C (72). The concentration was determined by BCA protein assay kit (Bio-Rad). The primary antibodies are SIX3 (Abcam, catalog no. AB172131), HESX1 (ABclonal, catalog no. A10696), PAX6 (Covance, catalog no. AB2237), and GAPDH (ABclonal, catalog no. AC002) (supplemental Table S3). The lysates were resolved by SDS-PAGE, and

Western blotting was carried out using horseradish peroxidase–conjugated IgG as a secondary antibody and enhanced chemiluminescence system for detection (ABclonal).

RNA-seq library construction

The RNA-seq library construction follows the method published previously (73). cDNA samples were sheared by ultrasonication on a Covaris S2 for 80 s with the following parameter settings: duty cycle = 10; cycles per burst = 200; intensity = 4. Samples were ethanol-precipitated and resuspended in Tris buffer for subsequent enzymatic modifications. End repair was carried out with T4 DNA polymerase and T4 polynucleotide kinase (New England Biolabs), and samples were purified with Ampure XP paramagnetic beads (Beckman Coulter). Blunt-end, 3′-phosphorylated products were 3′-adenylated with exo-Klenow fragment in the presence of dATPs (New England Biolabs), purified with Ampure XP beads, and ligated to sequencing adapters (Illumina) by T4 DNA ligase at 20 °C for 30 min. PCR amplification of library constructs was carried out with AccuPrime DNA polymerase (Invitrogen) for 13 cycles. Molarity and size distribution of sequencing libraries were assessed by HS-DNA microfluidic chips on the 2100 Bioanalyzer (Agilent). Sequencing was performed in 100-bp paired-end format on the Illumina HiSeq 2000.

RNA-Seq data preprocessing

Raw reads were mapped to the hg19 genome using TopHat2 version 2.0.4 program (74). We calculated FPKM as expression level using Cufflinks version 2.0.2 with default parameters (75). Then, we discarded genes that do not have FPKM >0.1 in at least one sample within all 12 time points during the hESC neural differentiation, and we next transformed expression levels to log-space by taking the $\log_2(\text{FPKM} + 1)$.

Weighted gene co-expression network analysis

A signed and weighted correlation network was constructed by first creating a matrix of pairwise correlations between all pairs of genes across the measured samples (26). Next, the adjacency matrix was constructed. Soft power parameter was estimated and is interpreted as a soft threshold of the correlation matrix. Based on the resulting adjacency matrix, we calculated the topological overlap, which is a robust and biologically meaningful measure of network interconnectedness (*i.e.* the strength of two genes' co-expression relationship with respect to all other genes in the network). Genes with highly similar co-expression relationships were grouped together by performing average linkage hierarchical clustering on the topological overlap. We used the Dynamic Hybrid Tree Cut algorithm to cut the hierarchical clustering tree, and we defined modules as branches from the tree cutting. We summarized the expression profile of each module by representing it as the first principal component (referred to as module eigengene). Modules whose eigengenes were highly correlated (correlation above 0.75) were merged.

Functional enrichment analysis

Functional enrichment of gene sets with different expression patterns was performed using the Database for Annotation, Visualization, and Integrated Discovery version 6.7 (76).

Author contributions—N. J. and J. D. J. H. conceived the study and supervised the project. N. J., J. D. J. H., Y. L., and R.W. designed the experiments. Y. L. and R. W. performed the experiments. R. W. and N. Q. analyzed the sequencing data. K. Z. helped with the CRISPR/Cas9 system. Y. L., R. W., K. T., G.P., J. D. J. H., and N. J. wrote the paper with the help of all other authors. All authors approved the final version of the manuscript.

Acknowledgments—We thank members of Jing laboratory for helpful discussions.

References

1. Kostovic, I., and Judas, M. (2006) Prolonged coexistence of transient and permanent circuitry elements in the developing cerebral cortex of fetuses and preterm infants. *Dev. Med. Child Neurol.* **48,** 388–393
2. Rakic, P. (2009) Evolution of the neocortex: a perspective from developmental biology. *Nat. Rev. Neurosci.* **10,** 724–735
3. Preuss, T. M., Cáceres, M., Oldham, M. C., and Geschwind, D. H. (2004) Human brain evolution: insights from microarrays. *Nat. Rev. Genet.* **5,** 850–860
4. Kriegstein, A., Noctor, S., and Martínez-Cerdeño, V. (2006) Patterns of neural stem and progenitor cell division may underlie evolutionary cortical expansion. *Nat. Rev. Neurosci.* **7,** 883–890
5. Arnold, S. J., and Robertson, E. J. (2009) Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. *Nat. Rev. Mol. Cell Biol.* **10,** 91–103
6. Tam, P. P., and Loebel, D. A. (2007) Gene function in mouse embryogenesis: get set for gastrulation. *Nat. Rev. Genet.* **8,** 368–381
7. Lu, C. C., Brennan, J., and Robertson, E. J. (2001) From fertilization to gastrulation: axis formation in the mouse embryo. *Curr. Opin. Genet. Dev.* **11,** 384–392
8. Weinstein, D. C., and Hemmati-Brivanlou, A. (1997) Neural induction in *Xenopus laevis*: evidence for the default model. *Curr. Opin. Neurobiol.* **7,** 7–12
9. Di-Gregorio, A., Sancho, M., Stuckey, D. W., Crompton, L. A., Godwin, J., Mishina, Y., and Rodriguez, T. A. (2007) BMP signalling inhibits premature neural differentiation in the mouse embryo. *Development* **134,** 3359–3369
10. Bibel, M., Richter, J., Schrenk, K., Tucker, K. L., Staiger, V., Korte, M., Goetz, M., and Barde, Y. A. (2004) Differentiation of mouse embryonic stem cells into a defined neuronal lineage. *Nat. Neurosci.* **7,** 1003–1009
11. Watanabe, K., Kamiya, D., Nishiyama, A., Katayama, T., Nozaki, S., Kawasaki, H., Watanabe, Y., Mizuseki, K., and Sasai, Y. (2005) Directed differentiation of telencephalic precursors from embryonic stem cells. *Nat. Neurosci.* **8,** 288–296
12. Zhang, S. C., Wernig, M., Duncan, I. D., Brüstle, O., and Thomson, J. A. (2001) *In vitro* differentiation of transplantable neural precursors from human embryonic stem cells. *Nat. Biotechnol.* **19,** 1129–1133
13. Chambers, S. M., Fasano, C. A., Papapetrou, E. P., Tomishima, M., Sadelain, M., and Studer, L. (2009) Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat. Biotechnol.* **27,** 275–280
14. Pankratz, M. T., Li, X. J., Lavaute, T. M., Lyons, E. A., Chen, X., and Zhang, S. C. (2007) Directed neural differentiation of human embryonic stem cells via an obligated primitive anterior stage. *Stem Cells* **25,** 1511–1520
15. Elkabetz, Y., Panagiotakos, G., Al Shamy, G., Socci, N. D., Tabar, V., and Studer, L. (2008) Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. *Genes Dev.* **22,** 152–165
16. Li, X.-J., Du, Z.-W., Zarnowska, E. D., Pankratz, M., Hansen, L. O., Pearce, R. A., and Zhang, S.-C. (2005) Specification of motoneurons from human embryonic stem cells. *Nat. Biotechnol.* **23,** 215–221
17. Cho, M. S., Hwang, D. Y., and Kim, D. W. (2008) Efficient derivation of functional dopaminergic neurons from human embryonic stem cells on a large scale. *Nat. Protoc.* **3,** 1888–1894

18. Hu, B. Y., Du, Z. W., Li, X. J., Ayala, M., and Zhang, S. C. (2009) Human oligodendrocytes from embryonic stem cells: conserved SHH signaling networks and divergent FGF effects. *Development* **136,** 1443–1452
19. Bissonnette, C. J., Lyass, L., Bhattacharyya, B. J., Belmadani, A., Miller, R. J., and Kessler, J. A. (2011) The controlled generation of functional basal forebrain cholinergic neurons from human embryonic stem cells. *Stem Cells* **29,** 802–811
20. van de Leemput, J., Boles, N. C., Kiehl, T. R., Corneo, B., Lederman, P., Menon, V., Lee, C., Martinez, R. A., Levi, B. P., Thompson, C. L., Yao, S., Kaykas, A., Temple, S., and Fasano, C. A. (2014) CORTECON: a temporal transcriptome analysis of *in vitro* human cerebral cortex development from human embryonic stem cells. *Neuron* **83,** 51–68
21. Yao, Z., Mich, J. K., Ku, S., Menon, V., Krostag, A.-R., Martinez, R. A., Furchtgott, L., Mulholland, H., Bort, S., Fuqua, M. A., Gregor, B. W., Hodge, R. D., Jayabalu, A., May, R. C., Melton, S., *et al.* (2017) A single-cell roadmap of lineage bifurcation in human ESC models of embryonic brain development. *Cell Stem Cell* **20,** 120–134
22. Huang, C. T., Tao, Y., Lu, J., Jones, J. R., Fowler, L., Weick, J. P., and Zhang, S.-C. (2016) Time-course gene expression profiling reveals a novel role of non-canonical WNT signaling during neural induction. *Sci. Rep.* **6,** 32600
23. Hu, B.-Y., and Zhang, S.-C. (2010) Directed differentiation of neural-stem cells and subtype-specific neurons from hESCs. *Methods Mol. Biol.* **636,** 123–137
24. Treutlein, B., Brownfield, D. G., Wu, A. R., Neff, N. F., Mantalas, G. L., Espinoza, F. H., Desai, T. J., Krasnow, M. A., and Quake, S. R. (2014) Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* **509,** 371–375
25. Langfelder, P., and Horvath, S. (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9,** 559
26. Xue, Z., Huang, K., Cai, C., Cai, L., Jiang, C. Y., Feng, Y., Liu, Z., Zeng, Q., Cheng, L., Sun, Y. E., Liu, J. Y., Horvath, S., and Fan, G. (2013) Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* **500,** 593–597
27. Luo, Y., Coskun, V., Liang, A., Yu, J., Cheng, L., Ge, W., Shi, Z., Zhang, K., Li, C., Cui, Y., Lin, H., Luo, D., Wang, J., Lin, C., Dai, Z., *et al.* (2015) Single-cell transcriptome analyses reveal signals to activate dormant neural stem cells. *Cell* **161,** 1175–1186
28. Cajal, M., Lawson, K. A., Hill, B., Moreau, A., Rao, J., Ross, A., Collignon, J., and Camus, A. (2012) Clonal and molecular analysis of the prospective anterior neural boundary in the mouse embryo. *Development* **139,** 423–436
29. Peng, G., Suo, S., Chen, J., Chen, W., Liu, C., Yu, F., Wang, R., Chen, S., Sun, N., Cui, G., Song, L., Tam, P. P., Han, J. D., and Jing, N. (2016) Spatial transcriptome for the molecular annotation of lineage fates and cell identity in mid-gastrula mouse embryo. *Dev. Cell* **36,** 681–697
30. Kamiya, D., Banno, S., Sasai, N., Ohgushi, M., Inomata, H., Watanabe, K., Kawada, M., Yakura, R., Kiyonari, H., Nakao, K., Jakt, L. M., Nishikawa, S., and Sasai, Y. (2011) Intrinsic transition of embryonic stem-cell differentiation into neural progenitors. *Nature* **470,** 503–509
31. Kang, H. J., Kawasawa, Y. I., Cheng, F., Zhu, Y., Xu, X., Li, M., Sousa, A. M., Pletikos, M., Meyer, K. A., Sedmak, G., Guennel, T., Shin, Y., Johnson, M. B., Krsnik, Z., Mayer, S., *et al.* (2011) Spatio-temporal transcriptome of the human brain. *Nature* **478,** 483–489
32. Walker, M. G., Volkmuth, W., Sprinzak, E., Hodgson, D., and Klingler, T. (1999) Prediction of gene function by genome-scale expression analysis: prostate cancer-associated genes. *Genome Res.* **9,** 1198–1203
33. Li, L., Liu, C., Biechele, S., Zhu, Q., Song, L., Lanner, F., Jing, N., and Rossant, J. (2013) Location of transient ectodermal progenitor potential in mouse development. *Development* **140,** 4533–4543
34. Zhu, Q., Song, L., Peng, G., Sun, N., Chen, J., Zhang, T., Sheng, N., Tang, W., Qian, C., Qiao, Y., Tang, K., Han, J. D., Li, J., and Jing, N. (2014) The transcription factor Pou3f1 promotes neural fate commitment via activation of neural lineage genes and inhibition of external signaling pathways. *Elife* **3,** e02224
35. Wilson, D., Charoensawan, V., Kummerfeld, S. K., and Teichmann, S. A. (2008) DBD-taxonomically broad transcription factor predictions: new content and functionality. *Nucleic Acids Res.* **36,** D88–D92
36. Fuxman Bass, J. I., Diallo, A., Nelson, J., Soto, J. M., Myers, C. L., and Walhout, A. J. (2013) Using networks to measure similarity between genes: association index selection. *Nat. Methods* **10,** 1169–1176
37. Parish, E. V., Mason, J. O., and Price, D. J. (2016) Expression of Barhl2 and its relationship with Pax6 expression in the forebrain of the mouse embryo. *BMC Neurosci.* **17,** 76
38. Yeung, J., Ha, T. J., Swanson, D. J., and Goldowitz, D. (2016) A novel and multivalent role of Pax6 in cerebellar development. *J. Neurosci.* **36,** 9057–9069
39. Liu, W., Lagutin, O., Swindell, E., Jamrich, M., and Oliver, G. (2010) Neuroretina specification in mouse embryos requires S1x3-mediated suppression of Wnt8b in the anterior neural plate. *J. Clin. Invest.* **120,** 3568–3577
40. Carlin, D., Sepich, D., Grover, V. K., Cooper, M. K., Solnica-Krezel, L., and Inbal, A. (2012) Six3 cooperates with Hedgehog signaling to specify ventral telencephalon by promoting early expression of Foxg1a and repressing Wnt signaling. *Development* **139,** 2614–2624
41. Zhang, X., Huang, C. T., Chen, J., Pankratz, M. T., Xi, J., Li, J., Yang, Y., Lavaute, T. M., Li, X. J., Ayala, M., Bondarenko, G. I., Du, Z. W., Jin, Y., Golos, T. G., and Zhang, S. C. (2010) Pax6 is a human neuroectoderm cell fate determinant. *Cell Stem Cell* **7,** 90–100
42. Yoo, Y. D., Huang, C. T., Zhang, X., Lavaute, T. M., and Zhang, S.-C. (2011) Fibroblast growth factor regulates human neuroectoderm specification through ERK1/2-PARP-1 pathway. *Stem Cells* **29,** 1975–1982
43. Chovsepian, A., Empl, L., Correa, D., and Bareyre, F. M. (2017) Heterotopic transcallosal projections are present throughout the mouse cortex. *Front. Cell. Neurosci.* **11,** 36
44. Strangward, P., Haley, M. J., Shaw, T. N., Schwartz, J. M., Greig, R., Mironov, A., de Souza, J. B., Cruickshank, S. M., Craig, A. G., Milner, D. A., Jr., Allan, S. M., and Couper, K. N. (2017) A quantitative brain map of experimental cerebral malaria pathology. *PLoS Pathog.* **13,** e1006267
45. Brito-Moreira, J., Lourenco, M. V., Oliveira, M. M., Ribeiro, F. C., Ledo, J. H., Diniz, L. P., Vital, J. F., Magdesian, M. H., Melo, H. M., Barros-Aragão, F., de Souza, J. M., Alves-Leon, S. V., Gomes, FCA, Clarke, J. R., Figueiredo, C. P., *et al.* (2017) Interaction of amyloid-β (Aβ) oligomers with neurexin 2α and neuroligin 1 mediates synapse damage and memory loss in mice. *J. Biol. Chem.* **292,** 7327–7337
46. Miller, J. A., Ding, S. L., Sunkin, S. M., Smith, K. A., Ng, L., Szafer, A., Ebbert, A., Riley, Z. L., Royall, J. J., Aiona, K., Arnold, J. M., Bennet, C., Bertagnolli, D., Brouner, K., Butler, S., *et al.* (2014) Transcriptional landscape of the prenatal human brain. *Nature* **508,** 199–206
47. Lippmann, E. S., Williams, C. E., Ruhl, D. A., Estevez-Silva, M. C., Chapman, E. R., Coon, J. J., and Ashton, R. S. (2015) Deterministic HOX patterning in human pluripotent stem cell-derived neuroectoderm. *Stem Cell Reports* **4,** 632–644
48. Gaspard, N., Bouschet, T., Hourez, R., Dimidschstein, J., Naeije, G., van den Ameele, J., Espuny-Camacho, I., Herpoel, A., Passante, L., Schiffmann, S. N., Gaillard, A., and Vanderhaeghen, P. (2008) An intrinsic mechanism of corticogenesis from embryonic stem cells. *Nature* **455,** 351–357
49. Johnson, M. A., Weick, J. P., Pearce, R. A., and Zhang, S. C. (2007) Functional neural development from human embryonic stem cells: accelerated synaptic activity via astrocyte coculture. *J. Neurosci.* **27,** 3069–3077
50. Kim, J. E., O'Sullivan, M. L., Sanchez, C. A., Hwang, M., Israel, M. A., Brennand, K., Deerinck, T. J., Goldstein, L. S., Gage, F. H., Ellisman, M. H., and Ghosh, A. (2011) Investigating synapse formation and function using human pluripotent stem cell-derived neurons. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 3005–3010
51. Shi, Y., Kirwan, P., Smith, J., Robinson, H. P., and Livesey, F. J. (2012) Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses. *Nat. Neurosci.* **15,** 477–486
52. Linker, C., and Stern, C. D. (2004) Neural induction requires BMP inhibition only as a late step, and involves signals other than FGF and Wnt antagonists. *Development* **131,** 5671–5681
53. Ozair, M. Z., Kintner, C., and Brivanlou, A. H. (2013) Neural induction and early patterning in vertebrates. *Wiley Interdiscip. Rev. Dev. Biol.* **2,** 479–498
54. Sansom, S. N., and Livesey, F. J. (2009) Gradients in the brain: the control of the development of form and function in the cerebral cortex. *Cold Spring Harb. Perspect. Biol.* **1,** a002519–a002519

55. Bhinge, A., Poschmann, J., Namboori, S. C., Tian, X., Jia Hui Loh, S., Traczyk, A., Prabhakar, S., and Stanton, L. W. (2014) MiR-135b is a direct PAX6 target and specifies human neuroectoderm by inhibiting TGF-β/BMP signaling. *EMBO J.* **33,** 1271–1283
56. Wallis, D. E., Roessler, E., Hehr, U., Nanni, L., Wiltshire, T., Richieri-Costa, A., Gillessen-Kaesbach, G., Zackai, E. H., Rommens, J., and Muenke, M. (1999) Mutations in the homeodomain of the human SIX3 gene cause holoprosencephaly. *Nat. Genet.* **22,** 196–198
57. Pasquier, L., Dubourg, C., Blayau, M., Lazaro, L., Le Marec, B., David, V., and Odent, S. (2000) A new mutation in the six-domain of SIX3 gene causes holoprosencephaly. *Eur. J. Hum. Genet.* **8,** 797–800
58. Ribeiro, L. A., El-Jaick, K. B., Muenke, M., and Richieri-Costa, A. (2006) SIX3 mutations with holoprosencephaly. *Am. J. Med. Genet.* **140,** 2577–2583
59. Solomon, B. D., Lacbawan, F., Jain, M., Domené, S., Roessler, E., Moore, C., Dobyns, W. B., and Muenke, M. (2009) A novel SIX3 mutation segregates with holoprosencephaly in a large family. *Am. J. Med. Genet. A* **149,** 919–925
60. Lacbawan, F., Solomon, B. D., Roessler, E., El-Jaick, K., Domene, S., Vélez, J. I., Zhou, N., Hadley, D., Balog, J. Z., Long, R., Fryer, A., Smith, W., Omar, S., McLean, S. D., Clarkson, K., *et al.* (2009) Clinical spectrum of SIX3-associated mutations in holoprosencephaly: correlation between genotype, phenotype and function. *J. Med. Genet.* **46,** 389–398
61. Hehr, U., Pineda-Alvarez, D. E., Uyanik, G., Hu, P., Zhou, N., Hehr, A., Schell-Apacik, C., Altus, C., Daumer-Haas, C., Meiner, A., Steuernagel, P., Roessler, E., Winkler, J., and Muenke, M. (2010) Heterozygous mutations in SIX3 and SHH are associated with schizencephaly and further expand the clinical spectrum of holoprosencephaly. *Hum. Genet.* **127,** 555–561
62. Dattani, M. T., Martinez-Barbera, J. P., Thomas, P. Q., Brickman, J. M., Gupta, R., Mårtensson, I. L., Toresson, H., Fox, M., Wales, J. K., Hindmarsh, P. C., Krauss, S., Beddington, R. S., and Robinson, I. C. (1998) Mutations in the homeobox gene HESX1/Hesx1 associated with septo-optic dysplasia in human and mouse. *Nat. Genet.* **19,** 125–133
63. Thomas, P. Q., Dattani, M. T., Brickman, J. M., McNay, D., Warne, G., Zacharin, M., Cameron, F., Hurst, J., Woods, K., Dunger, D., Stanhope, R., Forrest, S., Robinson, I. C., and Beddington, R. S. (2001) Heterozygous HESX1 mutations associated with isolated congenital pituitary hypoplasia and septo-optic dysplasia. *Hum. Mol. Genet.* **10,** 39–45
64. Tajima, T., Hattorri, T., Nakajima, T., Okuhara, K., Sato, K., Abe, S., Nakae, J., and Fujieda, K. (2003) Sporadic heterozygous frameshift mutation of HESX1 causing pituitary and optic nerve hypoplasia and combined pituitary hormone deficiency in a Japanese patient. *J. Clin. Endocrinol. Metab.* **88,** 45–50
65. Cohen, R. N., Cohen, L. E., Botero, D., Yu, C., Sagar, A., Jurkiewicz, M., and Radovick, S. (2003) Enhanced repression by HESX1 as a cause of hypopituitarism and septooptic dysplasia. *J. Clin. Endocrinol. Metab.* **88,** 4832–4839
66. Carvalho, L. R., Woods, K. S., Mendonca, B. B., Marcal, N., Zamparini, A. L., Stifani, S., Brickman, J. M., Arnhold, I. J., and Dattani, M. T. (2003) A homozygous mutation in HESX1 is associated with evolving hypopituitarism due to impaired repressor-corepressor interaction. *J. Clin. Invest.* **112,** 1192–1201
67. Sobrier, M. L., Maghnie, M., Vié-Luton, M. P., Secco, A., di Iorgi, N., Lorini, R., and Amselem, S. (2006) Novel HESX1 mutations associated with a life-threatening neonatal phenotype, pituitary aplasia, but normally located posterior pituitary and no optic nerve abnormalities. *J. Clin. Endocrinol Metab.* **91,** 4528–4536
68. McNay, D. E., Turton, J. P., Kelberman, D., Woods, K. S., Brauner, R., Papadimitriou, A., Keller, E., Keller, A., Haufs, N., Krude, H., Shalet, S. M., and Dattani, M. T. (2007) HESX1 mutations are an uncommon cause of septooptic dysplasia and hypopituitarism. *J. Clin. Endocrinol. Metab.* **92,** 691–697
69. Perrier, A. L., Tabar, V., Barberi, T., Rubio, M. E., Bruses, J., Topf, N., Harrison, N. L., and Studer, L. (2004) Derivation of midbrain dopamine neurons from human embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 12543–12548
70. Chen, J., Suo, S., Tam, P. P., Han, J. J., Peng, G., and Jing, N. (2017) Spatial transcriptomic analysis of cryosectioned tissue samples with Geo-seq. *Nat. Protoc.* **12,** 566–580
71. Wu, Y., Liang, D., Wang, Y., Bai, M., Tang, W., Bao, S., Yan, Z., Li, D., and Li, J. (2013) Correction of a genetic disease in mouse via use of CRISPR-Cas9. *Cell Stem Cell* **13,** 659–662
72. Sheng, N., Xie, Z., Wang, C., Bai, G., Zhang, K., Zhu, Q., Song, J., Guillemot, F., Chen, Y. G., Lin, A., and Jing, N. (2010) Retinoic acid regulates bone morphogenic protein signal duration by promoting the degradation of phosphorylated Smad1. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 18886–18891
73. Boroviak, T., Loos, R., Bertone, P., Smith, A., and Nichols, J. (2014) The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nat. Cell Biol.* **16,** 516–528
74. Trapnell, C., Pachter, L., and Salzberg, S. L. (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25,** 1105–1111
75. Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14,** R36
76. Huang da W., Sherman, B. T., and Lempicki, R. A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4,** 44–57